**National Quality Forum—Measure Testing (subcriteria 2a2, 2b2-2b7)**

**Measure Number** (*if previously endorsed*)**:** N/A

**Measure Title**: PointRight OnPoint-30 SNF Rehospitalizations

**Date of Submission**: 2/5/2014

**Type of Measure:**

|  |  |
| --- | --- |
| Composite – ***STOP – use composite testing form*** | Outcome (*including PRO-PM*) |
| Cost/resource | Process |
| Efficiency | Structure |

|  |
| --- |
| **Instructions**   * Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form. * **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.** * **For outcome and resource use measures**, section **2b4** also must be completed. * If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed. * Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed. * If you are unable to check a box, please highlight or shade the box for your response. * Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.*** Received permission from NQF staff to exceed 20 pages. * Contact NQF staff regarding questions. Check for resources at [Submitting Standards webpage](http://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx). |

|  |
| --- |
| **Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF’s evaluation criteria for testing.**  **2a2.** **Reliability testing** [**10**](#Note10) demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.  **2b2.** **Validity testing** [**11**](#Note11) demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.    **2b3.** Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [**12**](#Note12)  **AND**  If patient preference (e.g., informed decision making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [**13**](#Note13)  **2b4.** **For outcome measures and other measures when indicated** (e.g., resource use):   * **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors that influence the measured outcome (but not factors related to disparities in care or the quality of care) and are present at start of care; [**14**](#Note14)**,**[**15**](#Note15) and has demonstrated adequate discrimination and calibration   **OR**   * rationale/data support no risk adjustment/ stratification.   **2b5.** Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [**16**](#Note16) **differences in performance**;  **OR**  there is evidence of overall less-than-optimal performance.  **2b6.** **If multiple data sources/methods are specified, there is demonstration they produce comparable results**.  **2b7.** For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data),analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.  **Notes**  **10.** Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).  **11.** Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.  **12.** Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.  **13.** Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.  **14.** Risk factors that influence outcomes should not be specified as exclusions.  **15.** Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.  **16.** With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers. |

**1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE**

*Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.*

**1.1. What type of data was used for testing**? (*Check all the sources of data identified in the measure specifications and data used for testing the measure*. *Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

|  |  |
| --- | --- |
| **Measure Specified to Use Data From:**  **(*must be consistent with data sources entered in S.23*)** | **Measure Tested with Data From:** |
| abstracted from paper record | abstracted from paper record |
| administrative claims | administrative claims |
| clinical database/registry | clinical database/registry |
| abstracted from electronic health record | abstracted from electronic health record |
| eMeasure (HQMF) implemented in EHRs | eMeasure (HQMF) implemented in EHRs |
| other: Nursing Facility MDS 3.0 | other: Nursing Facility MDS 3.0 |

**1.2. If an existing dataset was used, identify the specific dataset** (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing facility MDS, facility health OASIS, clinical registry*).

Nursing Facility MDS 3.0

**1.3. What are the dates of the data used in testing**?

January 2011 through December 2012

**1.4. What levels of analysis** **were tested**? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

|  |  |
| --- | --- |
| **Measure Specified to Measure Performance of:**  **(*must be consistent with levels entered in item S.26*)** | **Measure Tested at Level of:** |
| individual clinician | individual clinician |
| group/practice | group/practice |
| hospital/facility/agency | hospital/facility/agency |
| health plan | health plan |
| other: Click here to describe | other: Click here to describe |

**1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)**? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Initial, testing was completed on 2,800 facilities to verify the calculations. Subsequent analyses were conducted on the national MDS database for all Medicare certified SNFs in the country. Presented below is summary information on the number and types of SNFs nationally in 2011 and 2012.

**TABLE 1. Number and Types of SNFs**

|  |  |  |
| --- | --- | --- |
|  | **2011** | **2012** |
| **Number of SNFs** | 15,693 | 15,690 |
| **For –Profit** | 10,758 | 10,832 |
| **Not-for-Profit** | 4,030 | 3,968 |
| **Government** | 905 | 890 |

\*Data from AHCA Quality Report 2013, based on CMS OSCAR data. Available at

<http://www.ahcancal.org/qualityreport/Documents/AHCA_2013QR_ONLINE.pdf>

**1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)**? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)   
Data reported from 2012 shows that there were 2,452,848 Medicare admissions and 798,513 Non-Medicare admissions to SNFs. The table below provides a breakdown of the descriptive characteristics of these patients.

**TABLE 2. Characteristics of Patients**

|  |  |  |
| --- | --- | --- |
|  | **Medicare Admissions** | **Non-Medicare Admissions** |
| **Age Category** |  |  |
| Under 65 | 10.6% | 26.3% |
| Age 65-84 | 53.7% | 43.1% |
| 85 and Older | 35.8% | 30.6% |
| Average Age | 78.8 | 74.4 |
| **Gender** |  |  |
| Male | 37.7% | 39.6% |
| Female | 62.3% | 60.4% |
| **Bed Mobility** |  |  |
| Independent | 4.8% | 9.8% |
| Supervision/Limited Assistance | 22.3% | 23.8% |
| Extensive Assistance/Total Dependence | 72.7% | 66.4% |
| **Transfer** |  |  |
| Independent | 2.5% | 6.7% |
| Supervision/Limited Assistance | 23.5% | 25.1% |
| Extensive Assistance/Total Dependence | 73.9% | 68.1% |
| **Eating** |  |  |
| Independent | 34.2% | 34.9% |
| Supervision/Limited Assistance | 47.7% | 47.2% |
| Extensive Assistance/Total Dependence | 18% | 17.9% |
| **Toilet Use** |  |  |
| Independent | 2.5% | 5.9% |
|  | **Medicare Admissions** | **Non-Medicare Admissions** |
| Supervision/Limited Assistance | 20.9% | 21.7% |
| Extensive Assistance/Total Dependence | 76.5% | 72.3% |
| **Bathing** |  |  |
| Independent | 1.3% | 2.1% |
| Supervision/Limited Assistance | 9.6% | 10.9% |
| Extensive Assistance/Total Dependence | 88.8% | 86.7% |
| **Race/Ethnicity** |  |  |
| American Indian | 0.0% | 0.0% |
| Asian | 2.7% | 3.4% |
| Black | 10.0% | 13.1% |
| Hispanic | 3.8% | 6.2% |
| White | 82.1% | 75.2% |
| Native Hawaiian or Pacific Islander | 0.2% | 0.4% |
| Unknown | 1.1% | 1.8% |
| **Common Active Diagnoses** |  |  |
| Anemia | 31.0% | 26.1% |
| Arteriosclerotic Heart Disease | 17.9% | 21.5% |
| Congestive Heart Failure | 22.8% | 17.1% |
| COPD | 25.2% | 21.5% |
| Depression | 32.3% | 33.4% |
| Diabetes | 34.4% | 33.3% |
| Hip Fracture | 7.1% | 5.0% |
| Hypertension | 75.4% | 71.2% |
| Osteoporosis | 44.9% | 11.1% |
| Stroke | 12.4% | 12.7% |
| **Special Treatment and Services** |  |  |
| Brain Injury | 0.1% | 0.9% |
| Hospice | 0.4% | 5.9% |
| IV Medication | 9.1% | 7.7% |
| Parenteral/IV Nutrition | 0.6% | 0.4% |
| Respite | 0.0% | 0.9% |
| Ventilator/Respirator | 0.4% | 0.7% |

\*Data from AHCA Quality Report 2013, based on MDS and OSCAR data. Available at

<http://www.ahcancal.org/qualityreport/Documents/AHCA_2013QR_ONLINE.pdf>

**1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below**.

N/A

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2a2. RELIABILITY TESTING**

***Note****: If accuracy/correctness (validity) of data elements was empirically tested*, *separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter “see section 2b2 for validity testing of data elements”; and skip 2a2.3 and 2a2.4.*

**2a2.1. What level of reliability testing was conducted**? (*may be one or both levels*)  
 **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)  
 **Performance measure score** (e.g., *signal-to-noise analysis*)  
  
**2a2.2. For each level checked above, describe the method of reliability testing and what it tests** (*describe the steps―do not just name a method; what type of error does it test; what statistical analysis was used*)  
We used parallel forms reliability by calculating several measures based on MDS 3.0 data submitted by over 2,800 SNFs directly to the research team and MDS 3.0 data from these same SNFs provided by CMS. We calculated the number of admission, tracking rate, observed rehospitalization rate and expected rehospitalization rate using both data sets and compared the results.

**2a2.3. For each level of testing checked above, what were the statistical results from reliability testing**? (e*.g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)  
The results of these reliability tests showed that in 206 cases (7%), numbers matched exactly on both the number of admissions and the tracking rate. In 1,869 cases (66%), the CMS data observed rate calculation minus the SNF data observed calculation was within 1%. In 2,652 cases (94%), the CMS data expected rate calculation minus the SNF data expected calculation was within 1%.

**2a2.4 What is your interpretation of the results in terms of demonstrating reliability**? (i*.e., what do the results mean and what are the norms for the test conducted?*)  
MDS 3.0 data received from CMS are reliable when compared to data gathered directly from participating SNFs. We assumed that data gathered directly from SNFs would be more accurate and complete because the facilities providing these data were paying for analytic services.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b2. VALIDITY TESTING**

**2b2.1. What level of validity testing was conducted**? (*may be one or both levels*)  
 **Critical data elements** (*data element validity must address ALL critical data elements*)

**Performance measure score**

**Empirical validity testing** **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

**2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests** (*describe the steps―do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

First, Medicare hospitalization claims were used to validate MDS 3.0 discharge assessments. Discharge records were categorized into four groups based on the values in the discharge status field: acute (if acute hospital), non-acute (if psychiatric hospital or MR/DD facility), death (if deceased), and other (if community, another SNF or swing bed, IRF, Hospice and other). Any hospitalization claims within ±3 days of a discharge assessment for the individual were identified and checked whether the discharge status could be verified by the Medicare claim. Verifying claims as acute (i.e. inpatient claims filed by general hospitals), outpatient and non-acute (i.e. inpatient claims filed by specialized hospitals) were grouped. Death was verified using date of death from the enrollment records. Discharge records matching to hospital claims and death dates were examined.

Second, the proportion of Medicare hospital claims that had an associated MDS 3.0 discharge assessment designated as being sent to the hospital from the SNF were estimated. The origin location of patients based on previous MDS 3.0 discharge assessments were also identified to ensure patients had not been discharged to a facility or other places after SNF admission. The 30 day rehospitalization rates were calculated using different data sources aggregating a binary variable indicating whether the patient was rehospitalized within 30 days of SNF admission. The rates were decomposed into three components: verified by other source, not rehospitalized from SNF and not-verified. If an individual had multiple hospitalizations, the earlier component trumped the later components.

The extent to which there was systematic error (related to facility characteristics) when MDS and claims data disagreed was explored. To examine this, the fraction of hospitalization events for SNFs that were identified by both MDS 3.0 and Medicare claims data was calculated. The hospitalization events that originated from SNFs and occurred within 30 days of SNF admission were included. The relationship of this variable with several SNF characteristics, including structural characteristics from OSCAR (size, occupancy rate, availability of staffing, deficiency score) and patient composition based on MDS 2010 were examined.

Finally, 30 day rehospitalization rates calculated at the SNF level based on acute discharge MDS 3.0 assessments with respect to rates based on Medicare hospitalization claims were plotted.

We also conducted construct validity testing. We hypothesized that facilities with low rehospitalization rates would correlate with other measures of quality such as CMS’s overall five star rating system, the staffing component of the five star rating system, the number of survey deficiencies cited by CMS during their annual onsite inspection, as well as AHCA’s quality award program based on the Baldridge program. We did not test a relationship with the five star quality measure component since nearly all of the quality measures relate to long stay residents and the rehospitalization measure applies to short stay residents. We did test the relationship between a facility’s rehospitalization rate and the short stay quality measure for pneumococcal vaccine (since infection is a leading cause of hospitalization and high vaccination rates also indicate a facility with a systematic process and philosophy of prevention).

We grouped facilities by their quality measures (e.g. five star rating, pneumococcal vaccination rates by quintile, and recipients of AHCA’s Baldrige based quality award at silver or gold level) and calculated the rehospitalization rates for each grouping and also conducted correlation tests. We hypothesized that facilities with higher five star rating would have a lower rehospitalization for overall five star rating, the survey deficiency component and staffing component of five star. We also hypothesized that facilities with higher rates of pneumococcal vaccination would have lower rehospitalization rates. Additionally, we hypothesized that silver or gold recipients of the ACHA Quality Award program (based on meeting Baldrige criteria) would have lower rehospitalization rates compared to non-recipients.

**2b2.3. What were the statistical results from validity testing**? (*e.g., correlation; t-test*)  
As shown in Table 3, 82.9% of MDS 3.0 discharge assessments indicating discharge location at an acute care hospital could be verified with inpatient claims data. An additional 3.7% of MDS 3.0 discharges could be verified with outpatient claims (indicating the event had been billed as an observation stay that probably lasted at least one night). Altogether, only 12.9% of MDS 3.0 discharges indicating acute hospitalization could not be verified with Medicare claims data. Since most MDS 3.0 discharge records indicating non-acute hospitalization had a corresponding acute hospital Medicare claim (63.3%), if both types of MDS 3.0 hospitalizations were combined, the percentage of MDS 3.0 discharges verified by an inpatient, outpatient or non-acute hospital claim is 87%. As shown in the lower panel of Table 3, results were similar when discharge assessments were restricted to just those occurring within 30 days of SNF admission. It should be noted that MDS 3.0 discharge records indicating that the patient was discharged dead were extremely accurate in comparison with the Medicare encounter record, which includes a discharge death date. This is a great improvement over the performance of the MDS 2.0 discharge record based upon published analyses from the last decade.   
**TABLE 3. Verifying MDS Discharge Records Using Inpatient and Outpatient Claims and Death Records**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Discharge code from MDS** | **Hospitalization from Claims Data** | | | **Death from Enrollment** | **Not Verified** | **Total Number of Discharges** |
| Acute | Out-Patient | Non-Acute |
| **All Discharge Records Between 1/11-11/11** | Acute | 82.9 | 3.7 | 0.5 | 0.0 | 12.9 | 404,122 |
| Non-Acute | 63.3 | 1.1 | 18.4 | 0.0 | 17.2 | 9,283 |
| Dead | 0.0 | 0.0 | 0.0 | 99.8 | 0.1 | 81,220 |
| Others | 7.9 | 1.1 | 0.1 | 0.0 | 90.9 | 801,284 |
| **All Discharge Records within 30 Days of SNF Admission** | Acute | 82.5 | 0.5 | 3.7 | 0.0 | 13.3 | 216,674 |
| Non-Acute | 59.2 | 20.7 | 0.8 | 0.0 | 19.3 | 4,760 |
| Dead | 0.0 | 0.0 | 0.0 | 99.8 | 0.2 | 33,803 |
| Others | 3.5 | 0.1 | 0.6 | 0.0 | 95.8 | 470,370 |

Table 4 below shows the 30 day rehospitalization rates calculated using the different data sources as well as the proportion of these that can be verified using the alternate data source (e.g. MDS discharge records vs. Medicare claims). Rehospitalization rates based upon the MDS, whether including non-acute events or not, are lower than those relying upon Medicare claims. At least part of this is attributable to the fact that the Medicare claim is truly a 30 day rehospitalization rate regardless of whether the patient had been discharged from the hospital or not; whereas, the MDS discharge refers only to transfers directly from the SNF that occurred within 30 days of admission from the hospital. For example, of the 19.77 acute hospitalization rate measured by the presence of a Medicare inpatient acute hospital claim, the rate would be 15.81 if measured only from an MDS discharge directly from the SNF. However, 3.35 of the 19.77 were hospitalizations that occurred before 30 days but AFTER the patient was discharged from the SNF to another location, often facility, meaning that the unexplained differential in the Medicare claims based rate and the MDS based rate is only .61, or about 3 percent. Adding all the other sources of Medicare claims to the pool (including outpatient observation stays), the 30 day rehospitalization rate is higher at 21.11, but the proportion unaccounted for is still very small. The bottom two rows begin with the MDS based measures and ask how frequently they are confirmed by Medicare claims. In this instance, whether we combine the acute and non-acute or look only at them independently, 2.16 to 2.23, or just under 12%, of the difference is unexplained.

**Table 4. 30 Day Rehospitalization Rate Based Upon Different Data Sources**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source** | **30 day Rehospitalization Rate** | **Rehospitalization Rate Decomposed by Verification from Other Source** | | |
|  |  | Verified | Not from SNF | Non-verified |
| **Acute Hospitalization (Medicare Inpatient Claims)** | 19.77 | 15.81 | 3.35 | 0.61 |
| **Any Hospitalization (Medicare Inpatient, Outpatient and Chronic Hospital Claims)** | 21.11 | 16.69 | 3.8 | 0.62 |
| **Acute Hospitalization (MDS Discharge)** | 18.37 | 16.21 |  | 2.16 |
| **Any Hospitalization (MDS Discharge)** | 18.77 | 16.54 |  | 2.23 |

As shown in Table 5 below, after excluding Medicare claims for those individuals with prior MDS records indicating discharge from the nursing facility to facility or another location, 93% of Medicare hospitalization claims taking place within 30 days of SNF admission could be verified with MDS discharge records. Another 1.5% had MDS discharge records indicating discharge to a non-acute location and 5.6% did not have an MDS discharge assessment. This suggests that, relative to Medicare claims data, the MDS discharge record is about 94% accurate.  
**Table 5. Verifying Medicare Hospitalization Claims Using MDS Discharge Records**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Type of Claims** | **Total** | **% of Claims Identified as not from SNF** | **% of Claims that are Identified as from SNF Verified by MDS Records** | | |
| Acute | Non-Acute | Not verified |
| **All Hospitalization Claims within 30 Days of SNF Admission** | Acute | 241,559 | 20.2 | 93.0 | 1.5 | 5.6 |
| Non-acute | 15,439 | 39.1 | 90.2 | 0.5 | 9.3 |
| Out patient | 7,173 | 63.8 | 46.7 | 38.2 | 15.0 |
| All | 264,171 | 22.5 | 92.2 | 1.9 | 5.9 |

In order to determine whether SNFs with certain characteristics were more or less likely to have submitted discharge records on their patients that were “errors” relative to the “gold standard” of Medicare claims, we calculated the percentage of all MDS discharge records reported that corresponded to a Medicare claim (either inpatient or outpatient). Table 6 below presents the marginal effect of a one unit change in the explanatory variable on the facility “accuracy” rate. The facility characteristics generally have no relationship to the measurement performance as can be seen by t-statistics smaller than 2.0. The only variables that are marginally significant are the proportion of minority patients in the facility and even these are very small effects.

**Table 6. Linear Regression of Fraction of Hospitalization Events Identified From Both MDS and Medicare Claims onto SNF Characteristics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Independent variables** | **Coefficient** | **t-statistics** | **95% Confidence Interval** | |
| Current Health Deficiencies | -0.051 | -1.14 | -0.14 | 0.038 |
| Best-Guess Total Beds in Facility | 0.01 | 3.27 | 0.004 | 0.017 |
| % Medicaid as Primary Payer | 0.025 | 1.76 | -0.004 | 0.054 |
| Part of a Chain | 0.543 | 1.33 | -0.28 | 1.366 |
| For-Profit | -0.519 | -1.27 | -1.337 | 0.299 |
| Hospital Based | 0.248 | 0.3 | -1.394 | 1.89 |
| Resident Acuity Index | -0.184 | -1.58 | -0.418 | 0.05 |
| Any Physician Extender FTEs | -0.031 | -0.12 | -0.53 | 0.469 |
| Ratio of RN to total nurse | 0.673 | 0.45 | -2.342 | 3.688 |
| Total direct Care Hours per Day per Resident | -0.043 | -0.35 | -0.293 | 0.206 |
| Weighted Deficiency (all) Score | -0.003 | -0.92 | -0.01 | 0.004 |
| Percent Occupancy | 0.033 | 2.25 | 0.004 | 0.062 |
| % of Admissions Classified "Low Care" | 0.012 | 0.27 | -0.08 | 0.104 |
| % of Admissions from Acute Hospital | 0.012 | 0.84 | -0.017 | 0.042 |
| % of Admissions Female | 0.029 | 1.55 | -0.009 | 0.066 |
| % of Admissions Black | -0.055 | -3.96 | -0.083 | -0.027 |
| % of Admissions Hispanic | -0.068 | -3.14 | -0.112 | -0.025 |
| # of Annual Admissions per Bed | -0.104 | -1.31 | -0.264 | 0.056 |
| # Hospitalizations per Resident Year | 0.422 | 1.08 | -0.362 | 1.205 |
| Mean RUGS (512) Value Across Residents | 6.417 | 3.32 | 2.531 | 10.303 |
| Mean Age at Assessment Across Residents | 0.083 | 1.94 | -0.003 | 0.168 |
| Constant | 65.177 | 11.73 | 54.008 | 76.346 |
| N | 13684 |  |  |  |
| R-squared | 0.0121 |  |  |  |
| Joint Test of Significance F( 21,48) | 7.11 |  |  |  |

Finally, as shown in Figure 1 below, we compared facility level 30 day rehospitalization rates using MDS and Medicare claims data.

**FIGURE 1. 30 Day Rehospitalization Rates Calculated from Medicare Claims and MDS**



With respect to the relationship of rehospitalization rate and five Star rating, we found a consistent inverse relationship between the rehospitalization rate and the overall five star rating, health inspection component of five star staff, the nurse staffing level (see Table 7.A, 7.B, and 7.C below). The correlation coefficient for each table were -0.15739, -0.13402, -0.17274 respectively; all p <0.0001.

|  |  |
| --- | --- |
| **Overall Five Star Rating** | **2nd Quarter Risk Adjusted Rehospitalization Rate** |
| 1 | 19.0 |
| 2 | 18.3 |
| 3 | 17.9 |
| 4 | 17.3 |
| 5 | 16.4 |

**TABLE 7.A. Average Rehospitalization Rate by Overall Five Star Rating**

**TABLE 7.B Average Rehospitalization Rate by CMS’s Health Inspections Five Star Rating**

|  |  |
| --- | --- |
| **Health Inspection Rating on Five Star** | **2nd Quarter Risk Adjusted Rehospitalization Rate** |
| 1 | 18.6 |
| 2 | 18.0 |
| 3 | 17.6 |
| 4 | 17.0 |
| 5 | 16.2 |

**TABLE 7.C Average Rehospitalization Rate by Nurse Staffing Five Star Rating**

|  |  |
| --- | --- |
| **Nurse Staffing Rating on Five Star** | **2nd Quarter Risk Adjusted Rehospitalization Rate** |
| 1 | 19.0 |
| 2 | 18.5 |
| 3 | 18.0 |
| 4 | 17.2 |
| 5 | 15.2 |

With respect to the short stay quality measure for pneumococcal vaccination rates; we found an inverse correlation, -0.15916 P<0.0001 with a facility’s rehospitalization rate. In other words, the higher the vaccination rate, the lower the rehospitalization rate.

With respect to the relationship between a facility’s rehospitalization rate and being a recipient of AHCA’ Baldridge based award, silver/gold recipients have significantly lower rehospitalization rates compared to non-AHCA member recipients (17.8 vs 18.3 in 2011 Q4, p<0.01). This difference persisted 18 months later in 2013 Q2 data (17.2 vs 17.7, p<0.01).

The measure also detects change over time and has a normal distribution as shown below.  
**FIGURE 2. National Risk Adjusted Rehospitalization Rates**

As hypothesized the rehospitalization measure was correlated with other measures of quality. This supports using the rehospitalization measure as a measure of a SNF’s quality of care.   
 **2b2.4. What is your interpretation of the results in terms of demonstrating validity**? (i*.e., what do the results mean and what are the norms for the test conducted?*)  
In summary, the results suggest that the MDS 3.0 and Medicare claims correspond over 90%, depending upon how one collapses the types of hospitalization. As importantly, our analyses of the facility factors related to non-correspondence between MDS and Medicare claims strongly suggest that what we observe is truly “random error”, suggesting that an MDS based measure is a good surrogate for a Medicare claims based measure. The MDS based measure has the added advantage of being much timelier, includes Medicare advantage and even non-Medicare patients and clearly includes hospitalizations that are billed as “Observation Days,” which would be missed with a Medicare inpatient claims based measure.

While the team did not specifically conduct analysis on the MDS the validity of this tool has been confirmed by previous analyses presented in peer reviewed literature (list of citations provided below). Notably, Saliba and Buchanan (2008) found that the MDS 3.0 is reliable and valid. Validity of the instrument was determined by comparing items to established gold standards or other related items and scales. MDS 3.0 cognitive, depression and behavioral items have a higher level of correlation with the comparison groups than did MDS 2.0 items. Eighty one percent of the nurses sampled also strongly agreed or agreed that the MDS 3.0 was clinically relevant and 89 percent strongly agreed or agreed that the MDS 3.0 items provide a more accurate report of the resident’s characteristics.   
  
Saliba , D., & Buchanan, J. (2008). Development & validation of a revised nursing facility assessment tool: MDS 3.0. *Rand Health Corporation*.

Saliba, D. & Buchanan, J. (2012). Making the investment count: Revision of the minimum data set for nursing homes, MDS 3.0. *J Am Med Dir Assoc*. 13(7), 602-610.

Saliba, D., Buchanan, J., Eldelen, M.O., Streim, J., Ouslander, J., Berlowitz, D., & Chodosh, J. (2012). MDS 3.0: Brief interview for mental status. *J Am Med Dir Assoc.* 13(7), 611-617.

Saliba, D., DeFilippo, S., Edelen, M.O., Kroenke, K., Buchanan, J., & Streim, J. (2012) Testing the PHQ-9 interview and observational versions (PHQ-9 OV) for MDS 3.0. *J Am Dir Assoc.* 13(7), 618-625.

Saliba, D., Jones, M., Streim, J., Ouslander, J., Berlowitz, D., & Buchanan, J. (2012) Overview of significant changes in the minimum data set for nursing facilities version 3.0. *J Am Dir Assoc*. 13(7), 595-601.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b3. EXCLUSIONS ANALYSIS**

**NA**  **no exclusions — *skip to section*** [***2b4***](#section2b4)

**2b3.1. Describe the method of testing exclusions and what it tests** (*describe the steps―do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)  
 We included all Medicare/Medicaid certified SNFs in the US that were present from January 2011 through September 2012, except 85 facilities which could not be stratified into our categories due to extreme changes in the number of admissions. The final sample included 15,546 SNFs.

Variation in the number of admissions from hospitals to a SNF from one time period to the next is expected to affect the SNF’s rehospitalization rates. The fewer the number of admissions, the more volatile these changes in rates will be. Due to this we have decided to exclude SNFs with fewer than 30 admissions from hospitals during any 12 month period from our rehospitalization rate reporting. (Note: while rates for the excluded facilities are not reported, admissions and rehospitalizations from these facilities are used to calculate the national rate used in the calculation of the adjusted rehospitalization rate).

To test this decision we stratified SNFs based on their average number of admissions over four 12 month periods (Jan 11 to Dec 11, Apr 11 to Mar 12, July 11 to June 12 and Oct 11 to Sept 12). SNFs were stratified into four groups based on their average number of admissions: 1) those with fewer than 30 admissions, 2) 30-50 admissions, 3) 50-100 admissions, and 4) >100 admissions from hospital each year. We then compared the average change in rehospitalization rates from 12 month period to 12 month period across the four groups.

**2b3.2. What were the statistical results from testing exclusions**? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)  
As expected, the average change in rates decreased as the number of admissions increased. This is shown in the table below.  
**TABLE 8. Change in Rates Quarter to Quarter Stratified by Number Admissions to a SNF**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Category** | **Observations** | **Variable** | **Mean** | **S.D.** | **Maximum** |
| **Average<30** | 1589 | |  | | --- | | adjust\_rate\_Q1 | | changeQ21 | | changeQ32 | | changeQ43 | | |  | | --- | | 0.1314 | | 0.0531 | | 0.0574 | | 0.0567 | | |  | | --- | | 0.1632 | | 0.0887 | | 0.1509 | | 0.1619 | | |  | | --- | | 2.7674 | | 1.0841 | | 4.3926 | | 4.4453 | |
| **Average>=30 & <50** | 1220 | |  | | --- | | adjust\_rate\_Q1 | | changeQ21 | | changeQ32 | | changeQ43 | | |  | | --- | | 0.1533 | | 0.0325 | | 0.0309 | | 0.0298 | | |  | | --- | | 0.0730 | | 0.0276 | | 0.0269 | | 0.0260 | | |  | | --- | | 0.5444 | | 0.2076 | | 0.2359 | | 0.1691 | |
| **Average>=50 & <100** | 2830 | |  | | --- | | adjust\_rate\_Q1 | | changeQ21 | | changeQ32 | | changeQ43 | | |  | | --- | | 0.1717 | | 0.0267 | | 0.0248 | | 0.0241 | | |  | | --- | | 0.0615 | | 0.0216 | | 0.0197 | | 0.0202 | | |  | | --- | | 0.4769 | | 0.1539 | | 0.1316 | | 0.1690 | |  | |
| **Category** | **Observations** | **Variable** | **Mean** | **S.D.** | **Maximum** |
| **Average>=100** | |  | | --- | | 9822 | | |  | | --- | | adjust\_rate\_Q1 | | changeQ21 | | changeQ32 | | changeQ43 | | |  | | --- | | 0.1862 | | 0.0160 | | 0.0153 | | 0.0149 | | |  | | --- | | 0.0519 | | 0.0132 | | 0.0126 | | 0.0126 | | |  | | --- | | 0.4528 | | 0.1132 | | 0.1133 | | 0.1056 | |

**2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results?** (*i.e., the value outweighs the burden of increased data collection and analysis.*  *Note:* ***If patient preference is an exclusion****, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)  
The results show the average change decreased as the number of admissions increased mostly below 30 admissions per year compared to those > than 30 admissions a year. This validates the decision to exclude SNFs with fewer than 30 admissions from hospitals during any 12 month period from our rehospitalization rate reporting.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES**  
***If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section*** [***2b5***](#section2b5)***.***

**2b4.1. What method of controlling for differences in case mix is used?**

**No risk adjustment or stratification**

**Statistical risk model with** 33 **risk factors**

**Stratification by** Click here to enter number of categories **risk categories**

**Other,** Click here to enter description

**2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities**.   
N/A

**2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors used in the statistical risk model or for stratification by risk** (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care and not related to disparities*)  
A bootstrap analysis as well as a stability analysis on the variables was conducted.

We performed a bootstrap analysis of the coefficients for PointRight OnPoint-30 in the following way: We began with a sample of 585,572 admissions to SNFs from acute care hospitals with admission dates in CY2011. Data were used if the SNF involved had a discharge assessment completion rate of 95% or higher. We calculated the coefficients of the PointRight OnPoint-30 logistic regression model on 1000 subsamples of 292,786 admissions. The distributions for each of the coefficients are displayed in the following table (Table 9) and compared with the coefficients used in the PointRight OnPoint-30 model, which was developed using a slightly different sample comprising 600,000 admissions to SNFs.

The PointRight OnPoint-30 model is based on the assumption that its independent variables rarely change between the day of admission and the assessment reference date of the first MDS assessment. While we cannot assess this directly we can look at the change from the first to the second PPS assessment of Medicare patients who remain in the facility long enough for two assessments. Typically this will be the change from day 7 to day 14 of a post-acute stay. This provides a rough estimate of variable stability. Table 10 shows the rates of change between assessments that were 7 days apart (N= 203,386). Note that only four variables show rates of change – usually in the direction of improvement – of greater than 10%. These variables are those for cognitive impairment, total bowel incontinence, two-person assist, and continued oxygen therapy. For these four variables the table shows the prevalence of 1s in the model building sample and the coefficient in the PointRight OnPoint-30 model. Considering all of the facts, it appears that facility-level estimates of expected readmission rates are unlikely to be affected greatly by variable instability. When the PointRight OnPoint-30 model is applied to data collected on the day of admission it will slightly overestimate the expected risk, because some patients with values of 1 on the least stable IVs will become zeroes by the day of the first MDS assessment.

**2b4.4. What were the statistical results of the analyses used to select risk factors?**Bootstrap:

Table 9 shows the difference between the PointRight OnPoint-30 coefficients and the mean coefficients from the bootstrap analysis, expressed as actual values, standard deviation (S.D.) and percentage. It is evident that only a few variables have more than 10% variation from the bootstrap mean; for those variables the absolute value and/or the number of standard deviations is clinically acceptable.

**TABLE 9. Pointright Onpoint-30 Coefficients Compared with Mean from Bootstrap Sampling**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable Type** | **Independent Variable** | **PointRight OnPoint-30 Coefficient** | **Bootstrap Mean** | **S.D.** | **Difference** | **Difference in S.D.** | **Difference in %** |
| **Intercept** | Intercept | -2.825 | -2.819 | 0.019 | -0.006 | -0.32 | 0.2% |
| **Type of Admission** | Medicare | 0.554 | 0.555 | 0.015 | 0.000 | -0.03 | -0.1% |
|  | Re-entry | 0.140 | 0.125 | 0.011 | 0.015 | 1.30 | 10.6% |
| **Demographics** | Male | 0.162 | 0.158 | 0.010 | 0.005 | 0.48 | 2.9% |
|  | Age Under 65 | 0.177 | 0.177 | 0.013 | 0.000 | 0.02 | 0.2% |
| **Diagnoses** | Anemia | 0.092 | 0.092 | 0.010 | 0.000 | 0.02 | 0.2% |
|  | Asthma | 0.103 | 0.105 | 0.011 | -0.002 | -0.16 | -1.7% |
|  | Diabetes Mellitus | 0.046 | 0.062 | 0.014 | -0.016 | -1.15 | -34.6% |
| **Variable Type** | **Independent Variable** | **PointRight OnPoint-30 Coefficient** | **Bootstrap Mean** | **S.D.** | **Difference** | **Difference in S.D.** | **Difference in %** |
| **Diagnoses** | Diabetic Foot Ulcer | 0.146 | 0.139 | 0.044 | 0.007 | 0.17 | 5.0% |
|  | Heart Failure | 0.200 | 0.206 | 0.012 | -0.006 | -0.51 | -3.0% |
|  | Internal Bleeding | 0.892 | 0.912 | 0.040 | -0.020 | -0.49 | -2.2% |
|  | Pressure Ulcer (Stage 2) | 0.167 | 0.181 | 0.016 | -0.014 | -0.86 | -8.2% |
|  | Pressure Ulcer (Stage 3) | 0.133 | 0.197 | 0.030 | -0.063 | -2.12 | -47.5% |
|  | Pressure Ulcer (Stage 4) | 0.157 | 0.146 | 0.037 | 0.011 | 0.29 | 6.8% |
|  | Pressure Ulcer (Unstageable) | 0.181 | 0.163 | 0.020 | 0.018 | 0.92 | 10.2% |
|  | Respiratory Failure | 0.116 | 0.163 | 0.025 | -0.047 | -1.86 | -40.6% |
|  | Septicemia | 0.089 | 0.121 | 0.029 | -0.032 | -1.09 | -35.7% |
|  | Vascular Ulcer | 0.186 | 0.181 | 0.027 | 0.006 | 0.21 | 3.0% |
|  | Viral Hepatitis | 0.402 | 0.310 | 0.049 | 0.092 | 1.87 | 22.8% |
| **Symptom** | Daily Pain | 0.061 | 0.054 | 0.017 | 0.007 | 0.40 | 11.1% |
| **Functional Status** | Bowel Incontinence (Total) | 0.185 | 0.176 | 0.011 | 0.009 | 0.77 | 4.7% |
|  | Cognition Not Intact | 0.333 | 0.331 | 0.011 | 0.001 | 0.14 | 0.4% |
|  | Eating Dependence | 0.472 | 0.430 | 0.017 | 0.042 | 2.48 | 8.9% |
|  | Two-Person Assist for Any ADL | 0.239 | 0.226 | 0.011 | 0.013 | 1.21 | 5.3% |
| **Treatments Continued from Hospital** | Cancer Chemotherapy | 0.600 | 0.595 | 0.050 | 0.005 | 0.10 | 0.8% |
|  | Dialysis | 0.604 | 0.606 | 0.021 | -0.002 | -0.09 | -0.3% |
|  | Insulin | 0.178 | 0.159 | 0.015 | 0.018 | 1.21 | 10.3% |
|  | IV Fluids or Meds | 0.188 | 0.179 | 0.017 | 0.009 | 0.52 | 4.7% |
|  | Ostomy Care | 0.326 | 0.349 | 0.026 | -0.023 | -0.87 | -6.9% |
|  | Oxygen | 0.340 | 0.346 | 0.012 | -0.007 | -0.56 | -2.0% |
|  | Radiation Therapy | 0.611 | 0.489 | 0.069 | 0.122 | 1.77 | 19.9% |
| **Variable Type** | **Independent Variable** | **PointRight OnPoint-30 Coefficient** | **Bootstrap Mean** | **S.D.** | **Difference** | **Difference in S.D.** | **Difference in %** |
| **Treatments Continued from Hospital** | Tracheostomy Care | 0.134 | 0.170 | 0.040 | -0.037 | -0.91 | -27.5% |
| **Mitigating Factors** | End-Stage Prognosis | -0.785 | -0.729 | 0.056 | -0.056 | -1.00 | 7.1% |
|  | Hospice Care | -1.509 | -1.423 | 0.098 | -0.086 | -0.87 | 5.7% |

Variable Stability:

**TABLE 10: Variable Stability between Two Assessments Seven Days Apart**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **% Changing from 0 to 1** | **% Changing from 1 to 0** | **% Unchanged** | **Prevalence of 1s in Validation Sample** | **Coefficient in Model** |
| Medicare | 0% | 0% | 100% |  |  |
| Re-entry | 1% | 1% | 99% |  |  |
| Male | 0% | 0% | 100% |  |  |
| Age Under 65 | 0% | 0% | 100% |  |  |
| Anemia | 2% | 2% | 98% |  |  |
| Asthma | 1% | 2% | 99% |  |  |
| Diabetes Mellitus | 1% | 1% | 99% |  |  |
| Diabetic Foot Ulcer | 0% | 0% | 100% |  |  |
| Heart Failure | 1% | 1% | 99% |  |  |
| Internal Bleeding | 0% | 0% | 100% |  |  |
| Pressure Ulcer Stage 2 | 0% | 2% | 100% |  |  |
| Pressure Ulcer Stage 3 | 0% | 0% | 100% |  |  |
| Pressure Ulcer Stage 4 | 0% | 0% | 100% |  |  |
| Pressure Ulcer Unstageable | 0% | 1% | 100% |  |  |
| Respiratory Failure | 0% | 1% | 100% |  |  |
| Septicemia | 0% | 1% | 100% |  |  |
| Vascular Ulcer | 0% | 0% | 100% |  |  |
| Viral Hepatitis | 0% | 0% | 100% |  |  |
| Daily Pain | 2% | 4% | 98% |  |  |
| **Variable** | **% Changing from 0 to 1** | **% Changing from 1 to 0** | **% Unchanged** | **Prevalence of 1s in Validation Sample** | **Coefficient in Model** |
| Bowel Incontinence (Total) | 7% | 9% | 93% | 49% | 0.185 |
| Cognition Not Intact | 4% | 8% | 96% | 66% | 0.333 |
| Eating Dependence | 1% | 1% | 99% |  |  |
| Two-Person Assist | 4% | 14% | 96% | 57% | 0.239 |
| Chemotherapy | 0% | 1% | 100% |  |  |
| Dialysis | 0% | 3% | 100% |  |  |
| Insulin | 1% | 2% | 99% |  |  |
| IV Fluids or Medications | 0% | 6% | 100% |  |  |
| Ostomy Care | 0% | 0% | 100% |  |  |
| Oxygen | 0% | 18% | 100% | 22% | 0.34 |
| Radiation Therapy | 0% | 0% | 100% |  |  |
| Tracheostomy Care | 0% | 1% | 100% |  |  |
| End-Stage Prognosis | 0% | 0% | 100% |  |  |
| Hospice Care | 0% | 0% | 100% |  |  |

**2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach** (*describe the steps―do not just name a method; what statistical analysis was used*)  
A clinical panel reviewed the entire MDS for skilled nursing facilities, identifying items that might be expected on clinical grounds to correlate with 30 day readmission risk, and that would be unlikely to change between the day of hospital discharge and the day of the first MDS assessment – which takes place by day 8 of the stay for all Medicare patients. Such items included demographics, chronic disease diagnoses,  treatments begun in the hospital with orders to be continued in the SNF, and functional status items that change slowly when they change at all, such as the patient’s needing two-person assistance for transferring and/or bed mobility.  These items were screened for significant univariate associations with the dependent variable (readmission to any acute care hospital directly from the SNF within 30 days of admission). This process yielded 39 candidate variables. A logistic regression formula was then estimated utilizing the 39 candidates; this was progressively refined into one that utilized 33 independent variables. The six remaining ones – PTSD, transfusions, tuberculosis, continuing radiation therapy, continuing ventilator status, and continuing suction did not add explained variance if added to a model that already included the 33 actually used.  With the exception of ventilator status and suction, the variables all had relatively low prevalence in the model-building sample. Ventilator status and suction were strongly associated with tracheostomy care, so it was not surprising that only one of the three variables was significant in the multivariate model that we ultimately selected for risk adjustment of readmission rates.

*Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below*.  
***If stratified, skip to*** [***2b4.9***](#question2b49)

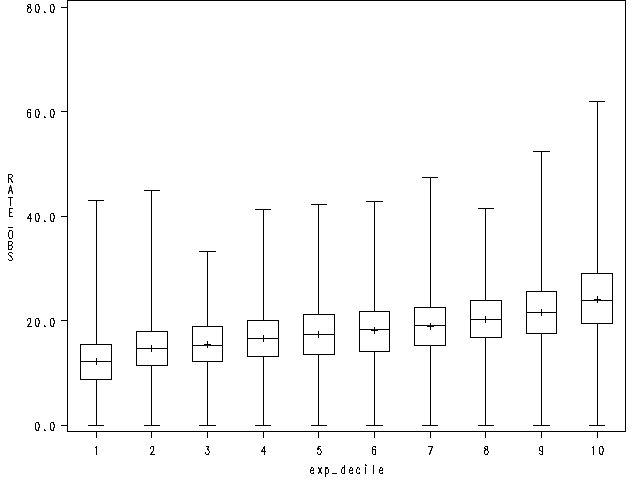
**2b4.6. Statistical Risk Model Discrimination Statistics** (*e.g., c-statistic, R-squared*)**:**

The c-statistic of the OnPoint-30 model is 0.669 with a 95% confidence interval (0:6666-0:6851). This means that there is 67% probability that a case (i.e. a person who gets rehospitalized) has higher predicted risk (i.e. higher estimated logit) than a non-case.

**2b4.7. Statistical Risk Model Calibration Statistics** (*e.g., Hosmer-Lemeshow statistic*):   
The p-value of the Hosmer-Lemeshow statistic for the OnPoint-30 model at the facility level is 0.85, so we accept the hypothesis of no discrepancy between Observed-Expected proportions, concluding that the logistic model is a good fit (well calibrated).

**2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves**:  
We grouped each facility into risk deciles based on their risk adjusted expected rehospitalization rate and then calculated the actual rate for each decile group. The “box and whisker” plot is shown in the figure below. As expected, the average actual rehospitalization rate increases steadily for each decile increase in expected rehospitalization indicating good calibration.

**FIGURE 3. CALIBRATION CURVE OF EXPECTED TO OBSERVED RATE**



**2b4.9. Results of Risk Stratification Analysis**:

N/A

**2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)?** (i*.e., what do the results mean and what are the norms for the test conducted*)  
N/A

**2b4.11.** **Optional Additional Testing for Risk Adjustment** (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

N/A

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE**

**2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified** (*describe the steps―do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

As shown in response to 2b3, the change in score across 1 quarter measured at three different quarters (Q1 to Q2; Q2 to Q3 and Q3 to Q4) were essentially the same size in magnitude for sample sizes greater than 30 but were much higher for sample sizes less than 30.

No tests of a clinically meaningful difference between providers or between observation periods were performed. The AHCA has set based on clinical input from membership and past experience that an improvement of 15% be considered clinically meaningful as part of AHCA’s quality initiative. Comparing facilities who are AHCA members’ rates from 2011 quarter 4 to 2013 quarter 2013 31.5% (2128 out of 6764) with reportable data achieved a 15% reduction in their rehospitalization rate.

**2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities?** (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

None performed

**2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities?** (i*.e., what do the results mean in terms of statistical and meaningful differences?*)

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS**

***If only one set of specifications, this section can be skipped.***

**Note***: This criterion is directed to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***If comparability is not demonstrated, the different specifications should be submitted as separate measures.***

**2b6.1. Describe the method of testing conducted to demonstrate comparability of performance scores for the same entities across the different data sources/specifications** (*describe the steps―do not just name a method; what statistical analysis was used*)  
N/A

**2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications?** (*e.g., correlation, rank order*)  
N/A

**2b6.3. What is your interpretation of the results in terms of demonstrating comparability of performance measure scores for the same entities across the different data sources/specifications?** (i*.e., what do the results mean and what are the norms for the test conducted*)  
N/A

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS**

**2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased** (*describe the steps―do not just name a method; what statistical analysis was used*)  
We determined the amount of missing data in the MDS 3.0 used to determine the numerator. If an MDS discharge assessment is not completed for individuals discharged a facility’s rates could be biased due to missing data. We specifically calculated the proportion of SNF admissions with missing discharge data.

This was accomplished as follows:

1. Identified all entries, including new admissions and re-entries. The entry date was determined using 2 variables: A1600 (entry date) and A0310F=01.
2. If an entry was accompanied by an admission assessment within 14 days of the entry date it was considered a new entry.
3. Determined whether an entry had a discharge record:

* If yes, then the entry was **complete**
* If not, then
  + If there was another entry record after the index entry, then the index entry is incomplete due to **missing discharge date**
  + If there was only one MDS entry record without any following MDS records:
    - If the entry date was within 14 days of 04/30/2012 (the latest date in the MDS 3.0 data received from AHCA), the reason for incompleteness could be due **to data truncation;**
    - If the entry date was a re-entry, and the date was within 120 days of 04/30/2012 (90 days would be the expected number, but an additional 30 days were added for increased tolerance), the reason of incompleteness could be due **to data truncation** (re-entry does not necessarily need an admission assessment or Medicare assessment)
    - If not above, the single entry record is problematic, and was assigned as “**only one record for an entry**”
  + If the last available MDS record was within 120 days of 04/30/2012, then the incompleteness could also be due to  **data truncation**
  + If not the above type, it was assigned the reason of incompleteness as “**all other incomplete**”

**2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data?** (*e.g.,**results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Results showed that the level of completeness is high, defined as over 95% of admissions having either a discharge assessment completed or another MDS data indicating that the person stayed in the facility, in most states as shown in Table 11. Based on these results, the decision was made to exclude facilities with greater than five percent missing data from the re-hospitalization rate analyses. In addition, facilities with between two and five percent missing data will be flagged in the reported re-hospitalization rates provided to the facility to allow them to improve their data completion rate. We also compared MDS data to claims data (see response to 2b2.3) and did not discover large amount of new cases not detected by MDS again supporting that missing data is infrequent for the majority of providers.

**TABLE 11. Distribution of MDS 3.0 Admission and Discharge Records and the Levels and Possible Types of “Missingness” by State**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **State** | **Number of**  **Admissions** | **Completed** | **Data Truncated** | **Missing Discharge** | **Only One MDS** | **All Other Incomplete** |
| **AK** | 1131 | 78.16% | 19.89% | 0.71% | 0.35% | 0.88% |
| **AL** | 81251 | 80.57% | 17.75% | 0.20% | 0.37% | 1.12% |
| **AR** | 52104 | 73.30% | 21.87% | 1.44% | 0.79% | 2.61% |
| **AZ** | 83897 | 88.41% | 9.80% | 0.20% | 0.69% | 0.90% |
| **CA** | 476657 | 83.35% | 14.10% | 0.32% | 0.46% | 1.40% |
| **CO** | 65252 | 83.32% | 15.23% | 0.20% | 0.46% | 0.78% |
| **CT** | 97330 | 82.84% | 15.49% | 0.28% | 0.44% | 0.95% |
| **DC** | 6213 | 74.51% | 22.71% | 0.56% | 0.66% | 1.56% |
| **DE** | 16916 | 83.18% | 15.80% | 0.18% | 0.31% | 0.54% |
| **FL** | 382468 | 85.14% | 13.49% | 0.23% | 0.35% | 0.79% |
| **GA** | 106487 | 79.43% | 19.92% | 0.15% | 0.20% | 0.30% |
| **HI** | 9552 | 80.31% | 17.50% | 0.13% | 0.34% | 1.73% |
| **IA** | 61458 | 75.69% | 23.40% | 0.18% | 0.21% | 0.52% |
| **ID** | 19087 | 85.48% | 13.77% | 0.07% | 0.36% | 0.32% |
| **IL** | 294768 | 82.15% | 16.26% | 0.25% | 0.38% | 0.96% |
| **IN** | 140985 | 81.18% | 18.30% | 0.10% | 0.12% | 0.30% |
| **KS** | 54171 | 77.57% | 21.25% | 0.16% | 0.35% | 0.68% |
| **State** | **Number of**  **Admissions** | **Completed** | **Data Truncated** | **Missing Discharge** | **Only One MDS** | **All Other Incomplete** |
| **KY** | 85170 | 81.47% | 18.07% | 0.08% | 0.09% | 0.29% |
| **LA** | 72386 | 75.19% | 23.91% | 0.31% | 0.16% | 0.43% |
| **MA** | 179603 | 84.06% | 14.72% | 0.16% | 0.43% | 0.63% |
| **MD** | 115055 | 83.47% | 14.26% | 0.29% | 0.51% | 1.46% |
| **ME** | 27241 | 84.68% | 14.78% | 0.11% | 0.29% | 0.14% |
| **MI** | 181933 | 83.83% | 14.38% | 0.30% | 0.37% | 1.11% |
| **MN** | 108197 | 84.02% | 15.63% | 0.01% | 0.12% | 0.22% |
| **MO** | 130969 | 79.05% | 18.85% | 0.45% | 0.47% | 1.18% |
| **MS** | 46460 | 76.98% | 22.17% | 0.41% | 0.18% | 0.26% |
| **MT** | 12995 | 79.57% | 19.48% | 0.17% | 0.31% | 0.47% |
| **NC** | 144286 | 82.56% | 16.97% | 0.12% | 0.13% | 0.22% |
| **ND** | 13377 | 74.37% | 25.04% | 0.07% | 0.19% | 0.33% |
| **NE** | 40226 | 79.88% | 19.43% | 0.13% | 0.17% | 0.40% |
| **NH** | 22624 | 81.10% | 18.29% | 0.07% | 0.18% | 0.36% |
| **NJ** | 215444 | 85.60% | 13.80% | 0.14% | 0.14% | 0.31% |
| **NM** | 21177 | 78.39% | 17.03% | 0.39% | 0.90% | 3.28% |
| **NV** | 31546 | 83.30% | 11.82% | 0.49% | 3.54% | 0.85% |
| **NY** | 356794 | 79.16% | 18.46% | 0.57% | 0.52% | 1.29% |
| **OH** | 334314 | 83.85% | 15.23% | 0.13% | 0.27% | 0.51% |
| **OK** | 55868 | 76.01% | 21.01% | 0.46% | 0.69% | 1.83% |
| **OR** | 47006 | 87.91% | 10.71% | 0.13% | 0.57% | 0.68% |
| **PA** | 327591 | 83.39% | 16.01% | 0.11% | 0.29% | 0.20% |
| **PR** | 833 | 97.84% | 1.32% | 0.00% | 0.72% | 0.12% |
| **RI** | 29537 | 81.38% | 16.47% | 0.33% | 0.67% | 1.16% |
| **SC** | 62031 | 80.89% | 17.38% | 0.26% | 0.49% | 0.99% |
| **SD** | 13294 | 73.34% | 26.40% | 0.04% | 0.15% | 0.08% |
| **TN** | 113035 | 80.36% | 18.14% | 0.23% | 0.43% | 0.84% |
| **TX** | 310019 | 73.74% | 20.58% | 0.89% | 0.80% | 3.98% |
| **UT** | 30841 | 84.53% | 12.28% | 0.27% | 0.74% | 2.18% |
| **VA** | 124513 | 83.63% | 15.12% | 0.15% | 0.27% | 0.84% |
| **VI** | 121 | 80.17% | 6.61% | 0.83% | 3.31% | 9.09% |
| **VT** | 10555 | 81.74% | 17.38% | 0.11% | 0.36% | 0.41% |
| **WA** | 90076 | 86.70% | 13.07% | 0.05% | 0.09% | 0.09% |
| **WI** | 105003 | 81.81% | 16.96% | 0.14% | 0.33% | 0.76% |
| **WV** | 33335 | 80.49% | 18.37% | 0.11% | 0.30% | 0.73% |
| **WY** | 5168 | 74.46% | 23.84% | 0.12% | 0.27% | 1.32% |

**2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased** due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Overall missing data is infrequent. The vast majority of providers had complete MDS data to calculate the measure, however, it is worthwhile to calculate the degree of missing data on the numerator and not report a facilities rates if they do not complete an MDS discharge assessment (the source for numerator information) at least 95% of the time.